

A Transcriptome-driven Analysis of Epithelial Brushings and Bronchial Biopsies to Define Asthma Phenotypes in U-BIOPRED

Chih-Hsi Scott Kuo^{1,2,3}, Stelios Pavlidis^{1,2,4}, Matthew Loza⁴, Fred Baribaud⁴, Anthony Rowe⁴, Ioannis Pandis^{1,2}, Uruj Hoda^{3,5}, Christos Rossios³, Ana Sousa⁶, Susan J. Wilson⁷, Peter Howarth⁷, Barbro Dahlen⁸, Sven-Erik Dahlen⁸, Pascal Chanez⁹, Dominick Shaw¹⁰, Norbert Krug¹¹, Thomas Sandström¹², Bertrand De Meulder¹³, Diane Lefaudeux¹³, Stephen Fowler¹⁴, Louise Fleming^{3,5}, Julie Corfield^{15,16}, Charles Auffray¹³, Peter J. Sterk¹⁷, Ratko Djukanovic⁷, Yike Guo^{1,2}, Ian M. Adcock^{3,5*}, and Kian Fan Chung^{3,5*}; on behalf of the U-BIOPRED Project Team[†]

¹Department of Computing, ²Data Science Institute, and ³Airways Disease Section, National Heart and Lung Institute, Imperial College London, London, United Kingdom; ⁴Janssen Research and Development, High Wycombe, United Kingdom; ⁵Biomedical Research Unit, Royal Brompton & Harefield National Health Service Trust, London, United Kingdom; ⁶Respiratory Therapeutic Unit, GlaxoSmithKline, Stockley Park, United Kingdom; ⁷Faculty of Medicine, Southampton University, Southampton, United Kingdom; ⁸Centre for Allergy Research, Karolinska Institute, Stockholm, Sweden; ⁹Université de la Méditerranée, Marseille, France; ¹⁰Centre for Respiratory Research, University of Nottingham, Nottingham, United Kingdom; ¹¹Fraunhofer Institute of Toxicology and Experimental Medicine, Hannover, Germany; ¹²Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden; ¹³European Institute for Systems Biology and Medicine, CNRS-ENS-UCBL, University of Lyon, Lyon, France; ¹⁴Centre for Respiratory Medicine and Allergy, University of Manchester, Manchester, United Kingdom; ¹⁵AstraZeneca R&D, Molndal, Sweden; ¹⁶Areteva R&D, Nottingham, United Kingdom; and ¹⁷Faculty of Medicine, University of Amsterdam, Amsterdam, the Netherlands

Abstract

Rationale: Asthma is a heterogeneous disease driven by diverse immunologic and inflammatory mechanisms.

Objectives: Using transcriptomic profiling of airway tissues, we sought to define the molecular phenotypes of severe asthma.

Methods: The transcriptome derived from bronchial biopsies and epithelial brushings of 107 subjects with moderate to severe asthma were annotated by gene set variation analysis using 42 gene signatures relevant to asthma, inflammation, and immune function. Topological data analysis of clinical and histologic data was performed to derive clusters, and the nearest shrunken centroid algorithm was used for signature refinement.

Measurements and Main Results: Nine gene set variation analysis signatures expressed in bronchial biopsies and airway epithelial brushings distinguished two distinct asthma subtypes associated with high expression of T-helper cell type 2 cytokines and lack of corticosteroid response (group 1 and group 3). Group 1 had the

highest submucosal eosinophils, as well as high fractional exhaled nitric oxide levels, exacerbation rates, and oral corticosteroid use, whereas group 3 patients showed the highest levels of sputum eosinophils and had a high body mass index. In contrast, group 2 and group 4 patients had an 86% and 64% probability, respectively, of having noneosinophilic inflammation. Using machine learning tools, we describe an inference scheme using the currently available inflammatory biomarkers sputum eosinophilia and fractional exhaled nitric oxide levels, along with oral corticosteroid use, that could predict the subtypes of gene expression within bronchial biopsies and epithelial cells with good sensitivity and specificity.

Conclusions: This analysis demonstrates the usefulness of a transcriptomics-driven approach to phenotyping that segments patients who may benefit the most from specific agents that target T-helper cell type 2-mediated inflammation and/or corticosteroid insensitivity.

Keywords: severe asthma; corticosteroid insensitivity; T-helper type 2; exhaled nitric oxide; gene set variation analysis

(Received in original form December 18, 2015; accepted in final form August 20, 2016)

*These authors contributed equally to this work.

†Other consortium project team members may be found in the online supplement.

Correspondence and requests for reprints should be addressed to Kian Fan Chung, M.D., D.Sc., Airways Disease Section, National Heart and Lung Institute, Imperial College London, Dovehouse Street, London SW3 6LY, UK. E-mail: f.chung@imperial.ac.uk

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Crit Care Med Vol 195, Iss 4, pp 443–455, Feb 15, 2017

Copyright © 2017 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201512-2452OC on August 31, 2016

Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the

Subject: Unbiased clustering of asthma using clinical features has derived phenotypes of limited pathobiological relevance. It is not known whether semibiased clustering of transcriptomic data will provide additional information on underlying asthma biology.

What This Study Adds to the

Field: We show that semibiased clustering of transcriptomic profiles derived from asthmatic airway samples produces biologically relevant clusters that have implications for directed or personalized therapy. We report subgroups of eosinophilic asthma driven by T-helper cell type 2 and corticosteroid insensitivity signatures that featured high submucosal eosinophils or sputum eosinophils. An inference scheme using currently available inflammatory biomarkers can be employed to predict the subtypes of gene expression.

Asthma is a heterogeneous disease that presents with varying degrees of inflammation in the airways in association with airway wall structural changes. Both infiltrating and resident airway structural cells participate in the inflammatory and remodeling processes of asthma, and various mechanisms have been proposed that could underlie the asthmatic inflammatory process (1). T-helper cell type 2 (Th2) activation characterized by the secretion of IL-4, IL-5, and IL-13 is a key mechanism of allergic asthma, with these cytokines being overexpressed in the bronchial submucosa of patients with

asthma (2, 3). Patients with Th2^{high} asthma are characterized by a greater degree of bronchial hyperresponsiveness, higher serum IgE levels, greater blood and airway eosinophilia, subepithelial fibrosis, and airway mucin gene expression (4).

Th1 and Th17 cells, as well as their products, are also important in asthma. Th1 cells are characterized by IFN- γ production, which plays a role in dealing with viral infections and in autoimmunity, and both Th1 cells and IFN- γ are overexpressed in patients with asthma (5–7). These have also been shown to be important in severe asthma (8, 9). Th17 cells express IL-17A, IL-17E, IL-17F, and IL-22, and they may mediate corticosteroid-resistant airway inflammation and airway hyperresponsiveness in mice (10). IL-17A and IL-17F have been localized in the airways of patients with severe asthma (11).

Oxidative stress is another important feature of asthma, particularly in patients with severe asthma, and it is associated with increased levels of reactive oxygen species (ROS) and compromised antioxidant responses (12, 13). ROS may mediate various features of severe asthma, such as the chronic inflammatory response, hypercontractility of the airways, and corticosteroid insensitivity (14–16). Finally, structural abnormalities can alter airway mechanics and contribute to bronchial hyperresponsiveness, and structural cells can also contribute to inflammatory processes through release of cytokines, chemokines, growth factors, and extracellular matrix elements (17–19).

Therefore, asthma is a syndrome with many potential endotypes, and defining these is an essential step toward providing personalized treatments (20, 21). Defining asthma phenotypes on the basis of clinical and physiologic parameters, including currently used inflammatory biomarkers, is limited in expressing the heterogeneity of underlying mechanisms (20, 22, 23). This

may not be surprising, considering the complexity of immunoinflammatory mechanisms and the diversity of cell types implicated in asthma (24–27). Deriving asthma clusters according to differentially expressed genes may be a better approach to capturing the diverse pathways of asthma pathobiology. It could lead to the identification of important targets for therapy and offer insights into achieving personalized medicine (20, 28, 29). This approach has been used to identify genes from sputum cells that linked pathway-based transcriptomic clusters to clinically important features of asthma (30).

Conventional gene set enrichment methods have limitations owing to the assumption of two classes (e.g., case vs. control) (31). Gene set variation analysis (GSVA) addresses this limitation by calculating gene set enrichment scores as a function of genes inside and outside a specifically defined gene set (32), and it can indicate functional activity across individual samples (33). Topological data analysis (TDA) that recognizes the invariant shape of complex datasets (34, 35) has been used to overcome the challenges that stochastic gene expression presents in conventional hierarchical or *k*-means clustering (36).

In the present study, we analyzed transcriptomic data derived from bronchial biopsies of participants with asthma from the U-BIOPRED (Unbiased Biomarkers in Prediction of Respiratory Disease Outcomes) project (37), and we performed supervised clustering using 42 gene sets associated with asthma and immune and/or inflammatory pathways by GSVA. We then explored the relationship between these GSVA signatures and their unique clinicohistopathologic features. This relationship was then further investigated in the transcriptome using epithelial brushings of the same participants with the aim of examining whether transcriptomic signatures may help define subsets of Th2^{high}

The U-BIOPRED project is supported through an Innovative Medicines Initiative joint undertaking under grant agreement 115010, resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007–2013) and European Federation of Pharmaceutical Industries and Associations companies' in-kind contributions (www.imi.europa.eu). C.-H.S.K. is supported by the Department of Thoracic Medicine, Linkou Chang-Gung Memorial Hospital, through a grant from Chang-Gung Medical Foundation, Taoyuan, Taiwan.

Author Contributions: C.-H.S.K., D.L., B.D.M., and S.P. performed the analysis; C.-H.S.K., S.P., M.L., F.B., A.R., C.A., I.P., and Y.G. designed the analytical approaches taken and analyzed the results; R.D., P.J.S., P.H., U.H., S.-E.D., P.C., D.S., B.D., N.K., T.S., and K.F.C. participated in the clinical characterization of the patients; S.J.W. and P.H. participated in the immunohistochemical studies; C.R. participated in experimental work and analysis; S.F., L.F., A.S., and J.C. were part of the data curation team; I.M.A., R.D., P.J.S., and K.F.C. conceived of and designed the study; and C.-H.S.K., I.M.A., and K.F.C. coordinated the data and drafted the manuscript. All authors read the final version of the manuscript.

The transcriptomic data have been deposited in the Gene Expression Omnibus database, <http://www.ncbi.nlm.nih.gov/geo> (accession no. GSE76225 for expression data of bronchial biopsies and GSE76226 for expression data of epithelial brushings).

eosinophilic and non-Th2 asthma from distinct airway compartments. Some of the results of these studies were reported previously in the form of an abstract (38).

Methods

Details of the analytical methods are provided in the online supplement.

Clinical Data

A total of 107 participants (Table 1) with moderate to severe asthma from the U-BIOPRED cohort underwent fiberoptic bronchoscopy for epithelial cell brushings and bronchial biopsies (39). Prebronchodilator spirometry, fractional exhaled nitric oxide (F_{ENO}), sputum differential cell count, skin-prick tests, serum total IgE, and differential blood count were measured (39). Oral corticosteroid dependence was assessed by determining the lowest maintenance dose of daily prednisolone. The study was approved by the ethics committees for each of the 16 clinical recruiting centers. All participants gave written and signed informed consent.

Immunohistochemistry for Submucosal Cells in Bronchial Biopsies

Samples from glycol methacrylate-embedded sections (2 μ m) were stained

with monoclonal antibodies against CD3, CD4, CD8, neutrophil elastase, and EG2. Cell counts were performed in a blinded fashion and expressed as positive cells per square millimeter.

Microarray Analysis of mRNA

Expression profiling was performed using the GeneChip® Human Genome U133 Plus 2.0 microarray (Affymetrix, Santa Clara, CA) with total RNA extracted from bronchial biopsies and epithelial brushings. RNA purity (RNA integrity number, >9.5) was measured using an Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA). Raw data were quality assessed and preprocessed by robust multiarray average normalization using the Almac Pipeline and Pre-processing Toolbox (Almac, Craigavon, UK). Probes of low expression were filtered according to robust multiarray signal values less than 5 and batch and/or technical effects were adjusted as covariates using a linear model.

GSVA

Forty-two gene sets containing 2,431 genes were manually curated: 6 for six key immune cell type-specific gene expression associated with their differentiation and activation; 6 for evolutionarily conserved transcriptional signatures for Th1 and CD8 memory T-cell differentiation; 5 of effector CD4⁺ T-cell differentiation for Th2, Th17, and regulatory T-cell subsets;

6 for chronic effects of oxidative stress in response to ozone from a mouse model and genes associated with chronic obstructive pulmonary disease (COPD) from human lung; 7 from peripheral blood mononuclear cells in four autoimmune diseases; 7 for asthma-specific mechanisms driven by T cells; and 5 from human lung biopsy, airway smooth muscle cells, and peripheral blood mononuclear cells in response to corticosteroid treatment (see Table E1 in the online supplement).

TDA and Machine Learning Algorithm

TDA was performed to explore clusters driven by gene signatures using Ayasdi Core software (Ayasdi, Inc., Menlo Park, CA), and the nearest shrunken centroid algorithm was used to further define the optimal number of classifiers. Details of the methods are provided in the online supplement.

Statistical Analysis

All datasets needed for this analysis were uploaded from the tranSMART system, an open-source knowledge management platform for sharing research data (40) that is supported by the European Translational Information and Knowledge Management Services ("eTRIKS") project. All categorical variables were analyzed using Fisher's exact test. Student's *t* test was used for continuous variables with normal distribution; otherwise, the Wilcoxon rank-sum test was used. A linear model for microarray data (Bioconductor R package limma) with a false discovery rate (FDR) correction was used for differential expression gene analysis, and recursive partitioning (party package in R; R Foundation for Statistical Computing, Vienna, Austria) was used for decision tree learning. An FDR less than 0.05 was considered statistically significant.

Results

Transcriptomic profiles from 91 bronchial biopsies and 99 epithelial brushings were obtained, of which 83 samples were paired from the same subject.

Differentially Expressed Genes and Gene Set Analysis

We performed differential expression analysis of individual genes and of gene sets comparing moderate asthma versus severe

Table 1. Demographic and Clinical Characteristics of 107 Subjects with Asthma

Variable	Data
Age, yr	45.9 $\pm$ 13.9
Female sex	59 (55.1%)
BMI, kg/m ²	28.7 $\pm$ 5.8
Severe asthma category	67 (62.6%)
Oral corticosteroid use	29 (27.1%)
Exacerbations, yr ⁻¹	2 (1–4)
Atopy	76 (71.0%)
Allergic rhinitis	53 (49.5%)
Eczema	42 (39.3%)
Nasal polyp	27 (25.2%)
FEV ₁ , % predicted	74.9 (62.3–95.1)
Total serum IgE, IU/ml	115.6 (43.0–316.0)
Blood eosinophils, %	3.1 (1.7–4.4)
Sputum eosinophils, %	1.5 (0.4–7.5)
Sputum neutrophils, %	52.5 (36.7–70.1)
F_{ENO} , ppb	31.5 (19.0–53.8)
Bronchial biopsy available	91 (85.0%)
Epithelial brushing available	99 (92.5%)

Definition of abbreviations: BMI = body mass index; F_{ENO} = fractional exhaled nitric oxide. Data are presented as n (%), mean $\pm$ SD, or median (interquartile range).

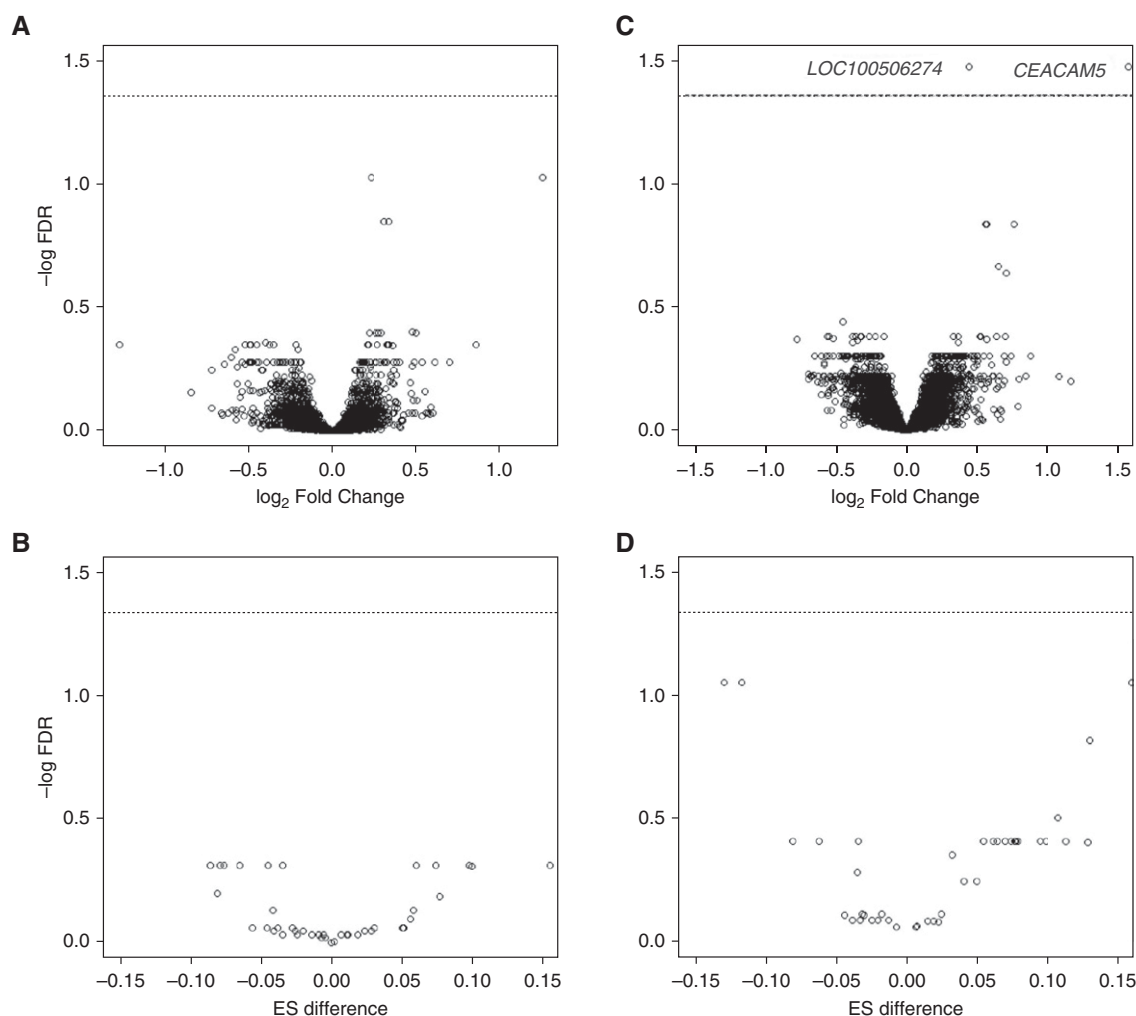


Figure 1. Differential expression analysis comparing moderate versus severe asthma from bronchial biopsies and epithelial brushing transcriptomic profiles on the basis of individual genes or gene set variation analysis. Given the false discovery rate (FDR) correction (dashed line), no differentially expressed genes (A) or gene sets (B) in epithelial brushing were noted. Two genes were marginally significant (C), and three gene sets (D) showed a trend toward significance in bronchial biopsies. CEACAM5 = carcinoembryonic antigen–related cell adhesion molecule 5; ES = enrichment score.

asthma in both epithelial brushings and bronchial biopsies. None of the individual genes or gene sets were significantly different between epithelial brushings from severe and nonsevere asthma after multiple testing correction for FDR (Figures 1A and 1B). In addition, we found only two genes with marginally significant expression and three GSVA gene sets with a trend toward significant expression when we compared expression profiles in bronchial biopsies of patients with severe asthma with those of patients with nonsevere asthma (Figures 1C and 1D). Therefore, the clinical classification of moderate versus severe asthma provides a limited framework for the identification of differential gene expression profiles.

There was a significant correlation between the enrichment score of the Th2 signature and the relative steroid insensitivity signature in both biopsy and brushings (Figure E1), together with a significant relationship between the oxidative stress signature and T-cell, Th1, and Th2 signatures (Figure E2), as well as between the oxidative stress signatures and relative corticosteroid insensitivity signatures in both biopsy and brushings (Figure E3). These correlations indicate potential mechanistic interactions between Th2 and Th1 genes with oxidative stress and relative corticosteroid insensitivity, which was defined by the signature of genes that are usually down-regulated by corticosteroids showing no change or being up-regulated.

In both biopsy and brushings, *CD44* as a constituent of corticosteroid insensitivity signature had the most extensive association and/or interaction with a subset of Th2 signature genes, including *CCL26*, *IL1R2*, and *CST2*. In addition, a highly correlated association between *NELFE* (within the corticosteroid insensitivity signature) and *ATP5J* (within the Th2 signature) was also seen in both biopsy and brushing compartments.

Clinical Clusters Driven by Bronchial Biopsy GSVA

We performed an analysis of the bronchial biopsy transcriptome using GSVA to investigate the immune

cell-specific characteristics and distinctive pathophysiologic processes in asthma. Unsupervised clustering of GSVA gene sets applied to biopsies using TDA identified a cluster characterized by high submucosal eosinophils (5.6 vs. 1.2/mm²; $P = 6.6 \times 10^{-5}$), high F_{ENO} (34.0 vs. 21.0 ppb; $P = 0.028$), and high oral corticosteroid use (41.0% vs. 15.4%; $P = 0.006$) (Figure 2 and Table 2). Comparing cluster A with cluster non-A, 26 of 42 gene sets were found to be differentially expressed ($P < 0.05$) (Figure 3). Similar clusters were obtained using hierarchical clustering (Figure 4).

Finding Signatures for Optimal Cluster Classification

To determine the GSVA signatures that best distinguished cluster A from cluster non-A, the shrunken centroid algorithm was applied to these 26 GSVA gene sets and produced a classification accuracy of 82.4% (Figure 5A). After feature reduction, nine GSVA gene sets of nonzero value that were expressed in the opposite direction in cluster A and cluster non-A were retained as centroid classifiers of discrimination (Figure 5B, Table 3). Receiver operating characteristic curve analysis demonstrated a high performance of this nine-gene-set signature model (area under the curve, 0.866; 95% confidence interval, 0.796–0.927; $P = 1.3 \times 10^{-9}$) in differentiating membership of cluster A, providing a sensitivity and specificity of 84.6% and 90.4%, respectively (Figure 5C).

Clusters from Epithelial Brushings and from Combined Epithelial Brushings and Bronchial Biopsies

To test whether these nine GSVA signatures were also able to discriminate clusters in bronchial brushings, a similar analysis was performed on the 99 epithelial brushing samples. The nine GSVA signatures from cluster A determined in bronchial biopsies were found in 31 (31.3%) of the subjects, as calculated using the nearest shrunken centroids of the signatures. This subgroup reproduced the clinical characteristics noted in biopsies in terms of significantly higher submucosal eosinophils (5.8 vs. 1.5/mm²; $P = 1.3 \times 10^{-5}$), higher F_{ENO} (51.5 vs. 21.0 ppb; $P = 4.2 \times 10^{-4}$), and higher oral corticosteroid use (40.7% vs. 17.9%; $P = 0.032$) (Table 4). In addition, higher eosinophil

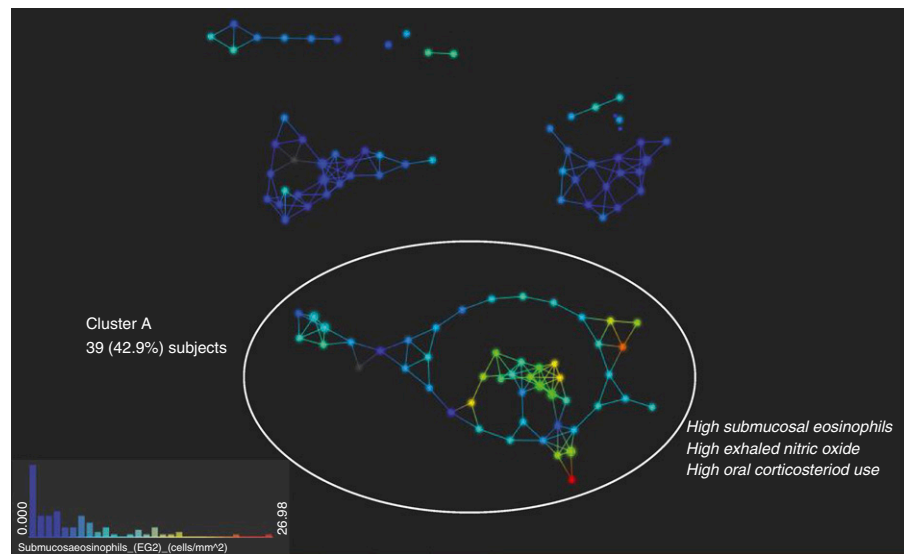


Figure 2. Topological network analysis of gene set variation analysis using 42 gene sets of relevance to the pathogenesis of asthma in bronchial biopsies revealing one large signature-driven cluster (cluster A) and four smaller distinct clusters. Cluster A was distinguished by higher submucosal eosinophil counts, higher levels of nitric oxide in exhaled breath, and higher oral corticosteroid use. The network is colored on the basis of the submucosal eosinophil count, with low cell count shown in blue and high cell count in red. The color key histogram (bottom left) shows the number of subjects with respect to the given cell count (x-axis) from low to high. Subjects with high submucosal eosinophil counts were noted to be highly enriched in cluster A as compared with the rest.

counts in blood and sputum, as well as higher numbers of submucosal CD3, CD4, and CD8 T cells in bronchial submucosa, were also noted.

The relative expression of these nine GSVA signatures in bronchial biopsies and epithelial brushings is shown in Table 5. Group 1 (19 subjects [23%]) expressed

Table 2. Clinical Characteristics of Cluster A versus Cluster Non-A

Variable	Cluster A (n = 39)	Cluster Non-A (n = 52)	P Value
Age, yr	48.4 ± 13.6	45.8 ± 13.1	0.355
Female sex	18 (46.2%)	31 (59.6%)	0.288
BMI, kg/m ²	28.1 ± 4.9	28.6 ± 5.8	0.661
Severe asthma	26 (66.7%)	30 (57.7%)	0.258
Oral corticosteroid use	16 (41.0%)	8 (15.4%)	0.006
Exacerbations, yr ⁻¹	2.5 (2.0–3.75)	2.0 (1.0–3.0)	0.139
Atopy	25 (64.1%)	40 (76.9%)	0.356
Allergic rhinitis	17 (43.6%)	26 (50.0%)	0.793
Eczema	15 (38.5%)	21 (40.4%)	1.000
Nasal polyp	14 (35.9%)	11 (21.2%)	0.092
FEV ₁ , % predicted	74.2 (59.8–91.6)	74.7 (62.3–94.6)	0.630
Total serum IgE, IU/ml	125.0 (41.0–231.5)	104.0 (44.0–364.5)	0.813
Blood eosinophils, %	3.2 (1.6–4.8)	2.7 (1.7–3.8)	0.523
Sputum eosinophils, %	3.9 (0.6–12.1)	1.1 (0.2–4.7)	0.156
Sputum neutrophils, %	53.2 (42.7–76.9)	52.4 (34.6–63.4)	0.421
F _{ENO} , ppb	34.0 (20.0–61.0)	21.0 (14.5–40.0)	0.028
Submucosal cells, mm ⁻²			
Eosinophils	5.6 (2.7–11.4)	1.2 (0–3.3)	6.6 × 10 ⁻⁵
Neutrophils	11.0 (6.6–18.1)	12.1 (6.8–18.4)	0.833
CD3 T cells	36.3 (22.6–50.0)	35.4 (16.6–49.5)	0.471
CD4 T cells	11.2 (5.9–18.9)	9.5 (4.8–15.8)	0.545
CD8 T cells	17.8 (9.5–30.4)	14.2 (4.9–27.7)	0.531

Definition of abbreviations: BMI = body mass index; F_{ENO} = fractional exhaled nitric oxide. Data are presented as n (%), mean ± SD, or median (interquartile range).

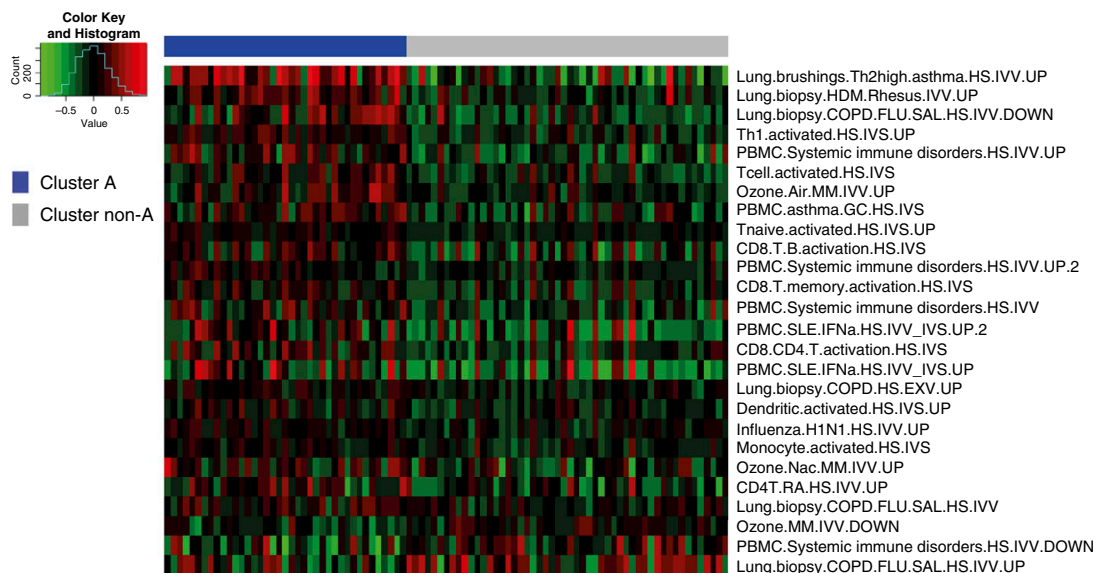


Figure 3. Heat map of differential gene set expression from bronchial biopsies between cluster A (blue bar) and cluster non-A (gray bar). High expression is denoted by red and low expression by green. Twenty-six of 42 gene sets used were significantly differentially expressed in cluster A, with 23 overexpressed and 3 underexpressed. COPD = chronic obstructive pulmonary disease; DOWN = down-regulated; EXV = ex vivo; FLU = fluticasone; GC = glucocorticoid; HDM = house dust mite; HS = *Homo sapiens*; IVS = in vitro soil-less; IVV = in vivo; MM = *Mus musculus*; PBMC = peripheral blood mononuclear cell; RA = rheumatoid arthritis; SAL = salmeterol; SLE = systemic lupus erythematosus; Th = T-helper cell; UP = up-regulated.

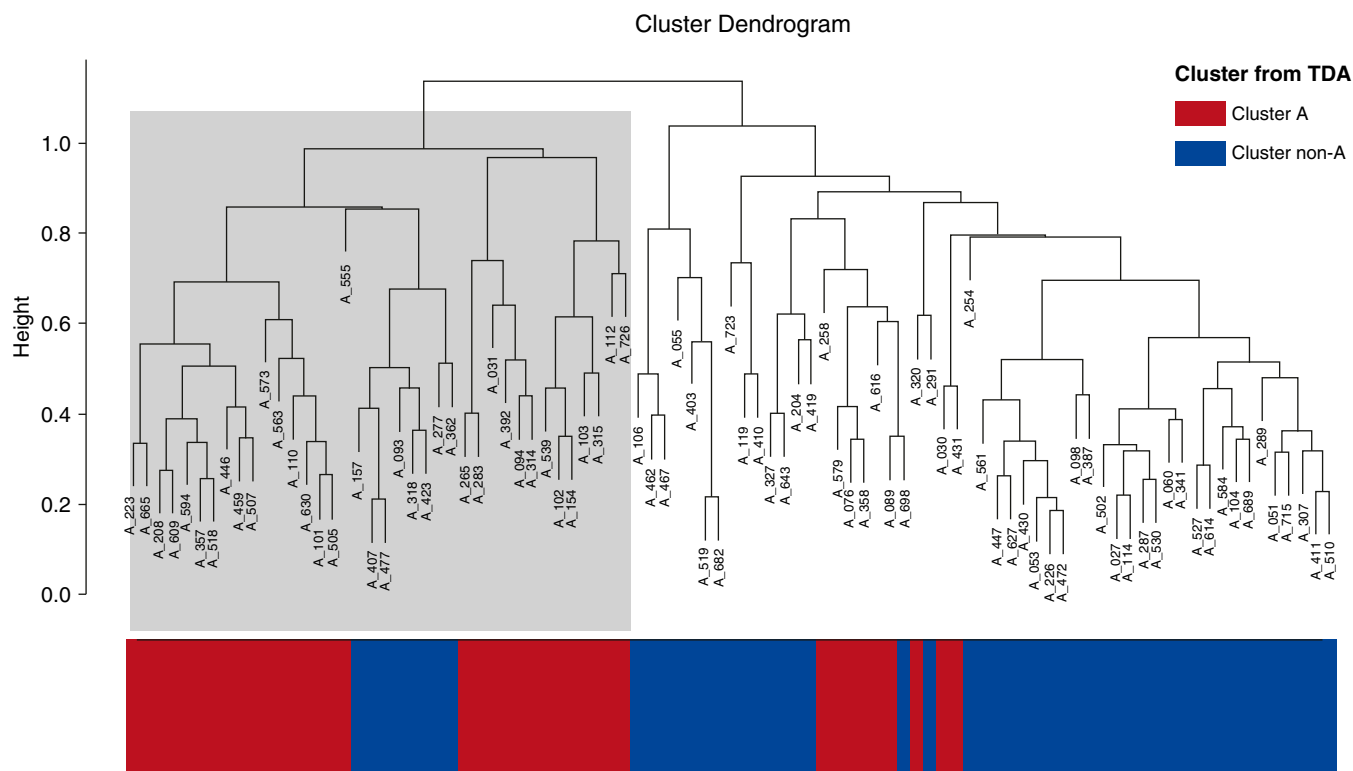


Figure 4. Hierarchical clustering showing 91 bronchial biopsies on 42 gene set variation analysis signatures. The color bar reveals the clustering result derived from topological data analysis (TDA) of the same dataset. Cluster A (red bar), identified by TDA, overlapped substantially with the subgroup of patients identified by hierarchical clustering (shaded in gray). The agreement levels of the two clustering methods measured by Rand index and adjusted Rand index were 0.7 and 0.4, respectively.

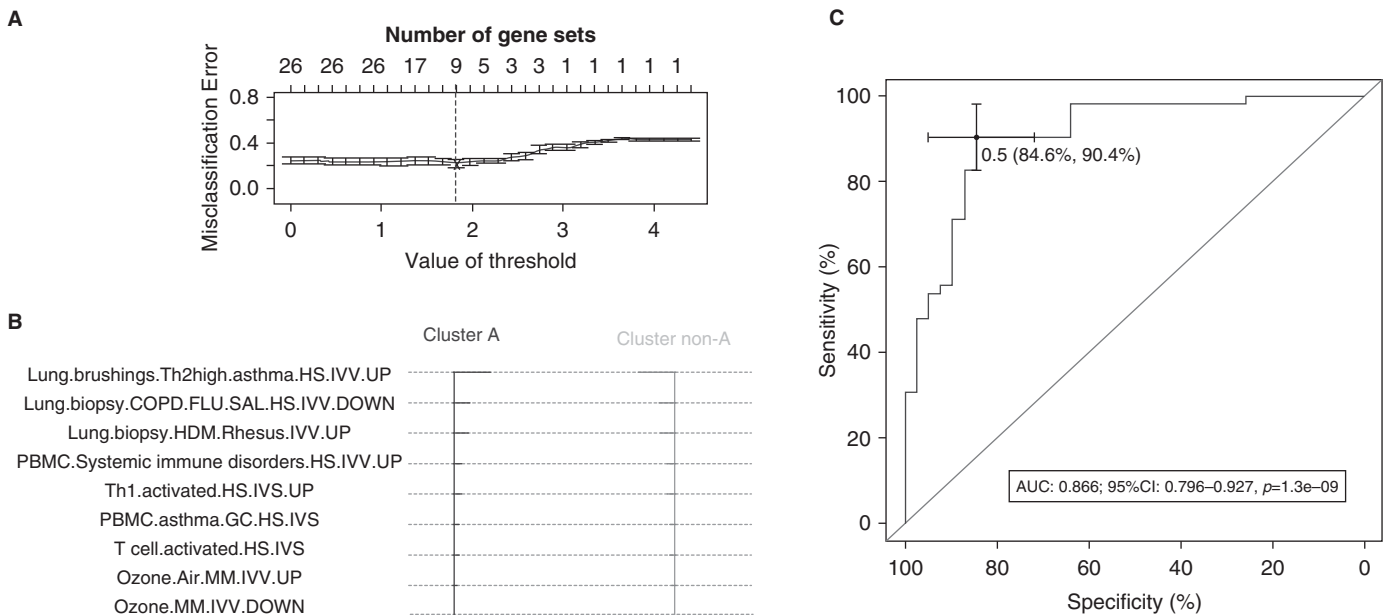


Figure 5. Machine learning algorithm using the shrunken centroid method on 26 differentially expressed gene sets in bronchial biopsies. (A) Training of classifiers for cluster A was evaluated for classification error using 10-fold cross-validation. A threshold of 1.82 (dashed line) was selected, which reduced classifiers to nine gene sets at an 82% cross-validated accuracy. Data are shown as mean $\pm$ SEM. (B) Centroid profile of the nine gene set signatures. The length of the centroid denotes the relative difference from the overall mean expression for each given gene set signature. From the top down, the centroids of cluster A and cluster non-A were ranked in decreasing amount. (C) Discrimination performance of the nine gene set signatures (area under the curve [AUC], 0.866; 95% confidence interval [CI], 0.796–0.927; $P=1.3 \times 10^{-9}$) based on the estimation of the cluster A probability derived from Gaussian density distribution. The probability cutoff at 0.5 gave a sensitivity and specificity of 84.6% and 90.4%, respectively, with the confidence intervals shown. COPD = chronic obstructive pulmonary disease; DOWN = down-regulated; FLU = fluticasone; GC = glucocorticoid; HDM = house dust mite; HS = *Homo sapiens*; IVS = *in vitro* soil-less; IVV = *in vivo*; MM = *Mus musculus*; PBMC = peripheral blood mononuclear cell; SAL = salmeterol; Th = T-helper cell; UP = up-regulated.

these nine signatures in both bronchial samples and epithelial brushing cells; group 2 (17 patients [20%]) expressed them in bronchial biopsies only; group 3 (8 patients [10%]) expressed them in epithelial brushing cells alone; and group 4 (39 patients [47%]) expressed them in neither compartment.

Compared with group 4 participants, two subtypes of patients with eosinophilic inflammation and relative steroid insensitivity were noted (groups 1 and 3) (Table 6). Group 1 had the highest submucosal eosinophils (6.1 vs. 1.0/mm²; $P=7.64 \times 10^{-6}$), CD3 T cells (45.3 vs. 33.3/mm²; $P=0.048$), CD8 T cells (26.4 vs. 13.9/mm²; $P=0.020$), F_{ENO} (56.5 vs. 19.0 ppb; $P=0.001$), acute exacerbation rate (3.0 vs. 1.5 times/yr; $P=0.022$), oral corticosteroid use (47.7% vs. 12.8%; $P=0.010$), and intermediate to high sputum eosinophils (8.2% vs. 0.6%; $P=0.017$). In contrast, group 3 participants possessed the highest levels of sputum eosinophils (15.8% vs. 0.6%; $P=0.014$), high body mass index (33.1 vs. 28.0 kg/m²; $P=0.010$), and a trend toward intermediate to high F_{ENO} (35.5 vs. 19.0 ppb; $P=0.058$).

Levels of gene expression of IL-33 and thymic stromal lymphopoietin (TSLP), but not of IL-25, were relatively increased in cluster A compared with cluster non-A and in group 1 compared with groups 3 and 4 (for IL-33) and with group 4 (for TSLP) in bronchial biopsies, but not in epithelial brushings (Figure E4).

Inference of Phenotype from Noninvasive Clinical Measurements

The clinical features and noninvasive biomarkers of the patients classified into the four groups were used to build an inference tree framework. F_{ENO} greater than 55 ppb (Figure 6, node 7, 58% for group 1), sputum eosinophils greater than 4.5% (Figure 6, node 6, 42% for group 1), or oral corticosteroid dependency (Figure 6, node 4, 36% for group 1) predicted subjects as the highest eosinophilic inflammation and steroid insensitivity subtype (group 1), with an 84% sensitivity and 72% specificity. Patients with F_{ENO} less than or equal to 55 ppb, sputum eosinophils less than or equal to 4.5%, and not dependent on oral corticosteroids had an 86% probability of having noneosinophilic inflammation

(Figure 6, node 5, 22% for group 2 and 64% for group 4). Clustering analysis based on the three top gene sets in the nine-gene-set GSVA signature associated with these clinical traits confirmed the molecular association of group 1 subjects with high Th2 activation and corticosteroid insensitivity in bronchial biopsy and epithelial brushing samples (Figure 7).

Discussion

We clustered gene expression data from bronchial biopsies and epithelial brushings obtained from patients with moderate to severe asthma to define clinical phenotypes of asthma. Direct comparison of the transcriptome obtained from subjects with severe asthma to moderately severe asthma led to the identification of few differentially expressed genes in bronchial biopsies and epithelial brushings using the FDR. We then examined their gene expression profiles on the basis of important disease drivers as described by specific gene sets and derived phenotypes from the gene set clusters. A collection of nine gene

Table 3. Nine Gene Set Variation Analysis Signatures Differentiating Cluster A from the Rest

GSVA Signature	ES Difference	P Value	FDR
Lung.brushings.Th2high.asthma.HS.IVV.UP	0.578	6.0×10^{-10}	2.5×10^{-8}
Lung.biopsy.COPD.FLU.SAL.HS.IVV.DOWN	0.308	5.3×10^{-6}	7.4×10^{-5}
Lung.biopsy.HDM.Rhesus.IVV.UP	0.275	7.8×10^{-7}	1.6×10^{-5}
PBMC.Systemic immune disorders.HS.IVV.UP	0.203	0.001	0.005
Th1.activated.HS.IVS.UP.2	0.199	7.2×10^{-6}	7.6×10^{-5}
PBMC.asthma.GC.HS.IVS	0.234	8.4×10^{-5}	7.1×10^{-4}
Tcell.activated.HS.IVS	0.192	1.5×10^{-4}	0.001
Ozone.Air.MM.IVV.Up	0.157	8.9×10^{-4}	0.004
Ozone.MM.IVV.Down	-0.159	2.1×10^{-4}	0.001

Definition of abbreviations: COPD = chronic obstructive pulmonary disease; DOWN = down-regulated; ES = enrichment score; FDR = false discovery rate; FLU = fluticasone; GC = glucocorticoid; GSVA = gene set variation analysis; HDM = house dust mite; HS = *Homo sapiens*; IVS = *in vitro* soil-less; IVV = *in vivo*; MM = *Mus musculus*; PBMC = peripheral blood mononuclear cell; SAL = salmeterol; Th = T-helper cell; UP = up-regulated.

set signatures applied to genes expressed in bronchial biopsies and airway epithelial brushings led to identification of two subtypes of patients with eosinophilic inflammation and relative corticosteroid insensitivity. Using machine learning tools, we found that currently available inflammatory biomarkers such as sputum eosinophilia and FE_{NO} levels could be used to predict the subtypes of patients with asthma described by gene expression profiling. This approach could ultimately define patients who may benefit the

most from specific molecular agents that target Th2-mediated inflammation and/or relative corticosteroid insensitivity. It also defines the importance of the site of expression of these nine gene set signatures in either biopsies or brushings, or in both, in determining the phenotype.

It is now generally accepted that asthma is a heterogeneous disease with distinct phenotypes. Our analysis confirmed the presence of a previously described Th2^{high} cluster derived from an analysis of Th2

gene expression in epithelial brushings that is indicative of Th2 activation (4, 41) and that was defined by 2 of the 42 GSVA gene sets we used. This signature has been linked to other features of asthma, particularly corticosteroid responsiveness. The remaining signatures we used reflected the heterogeneous nature of asthma and represent many of the mechanisms that have been proposed to drive the asthmatic disease process in addition to the Th2 pathway.

Interestingly, GSVA gene sets associated with steroid responsiveness accounted for the second and sixth most important signatures that were significantly up-regulated, when they are expected to be down-regulated by corticosteroids in patients with nonsevere asthma. In addition, a moderate correlation was noted between Th2 and corticosteroid insensitivity signatures (Figure E1), reflecting the clinical problem of clinical insensitivity to corticosteroid treatment in a subgroup of patients with severe asthma. Further analysis of the genes that were associated in the corticosteroid insensitivity and Th2 signatures in the biopsies and brushings showed an association between *CD44* and the Th2-associated genes *CCL26*, *IL1R2*, and *CST2*, indicating potential underlying mechanisms. The patients in group 1, where the corticosteroid insensitivity signatures were most highly expressed in the airway epithelium and bronchial biopsy, were receiving the highest doses of oral corticosteroids. Therapeutic strategies other than corticosteroid dose escalation should be considered for these patients.

Relative corticosteroid insensitivity was defined at the molecular level by detecting the signature of genes usually down-regulated by corticosteroids as being up-regulated. However, such a definition assumes that the patient is adherent to corticosteroid therapy, particularly oral therapy. The patients in the U-BIOPRED project have been followed for at least 6 months in a specialist severe asthma clinic, during which time their degree of compliance has been deemed to be satisfactory. Our observation that group 1, where the expression of the relative corticosteroid insensitivity signature together with the Th2^{high} and oxidative stress signatures were found in both bronchial biopsies and brushings, was associated with a

Table 4. Clinical Characteristics of the Nine-Gene-Set Signature Expressed in Epithelial Brushings

Variables	Expressed (n = 31)	Not Expressed (n = 68)	P Value
Age, yr	48.0 ± 14.8	44.4 ± 14.0	0.253
Female sex	12 (38.7%)	40 (58.8%)	0.101
BMI, kg/m ²	28.8 ± 5.4	29.0 ± 6.3	0.895
Severe asthma	23 (74.2%)	40 (58.8%)	0.211
Oral corticosteroid use	11 (40.7%)	10 (17.9%)	0.032
Exacerbations, yr ⁻¹	3.0 (2.0–4.0)	2.0 (1.0–3.0)	0.025
Atopy	17 (63.0%)	40 (75.0%)	0.419
Allergic rhinitis	19 (61.3%)	30 (44.1%)	0.171
Eczema	10 (32.3%)	32 (47.1%)	0.245
Nasal polyp	10 (32.3%)	15 (22.1%)	0.404
FEV ₁ , % predicted	74.2 (57.3–95.8)	87.9 (68.5–95.5)	0.264
Total serum IgE, IU/ml	184.0 (44.0–373.0)	87.0 (39.6–268.5)	0.067
Blood eosinophils, %	3.5 (2.7–7.4)	2.3 (1.5–3.8)	0.008
Sputum eosinophils, %	12.0 (2.6–26.2)	0.6 (0–1.7)	5.2×10^{-4}
Sputum neutrophils, %	53.2 (42.7–76.9)	52.4 (34.6–63.4)	0.786
FE _{NO} , ppb	51.5 (20.8–75.8)	21.0 (13.3–37.8)	4.2×10^{-4}
Submucosal cells, mm ⁻²			
Eosinophils	5.8 (3.0–12.5)	1.5 (0–4.8)	1.3×10^{-5}
Neutrophils	11.5 (4.6–17.1)	12.1 (6.8–19.2)	0.780
CD3 T cells	44.7 (32.2–61.7)	24.4 (14.8–46.8)	0.003
CD4 T cells	13.8 (7.9–23.7)	8.3 (3.2–14.6)	0.016
CD8 T cells	24.9 (15.7–37.6)	10.8 (4.2–20.2)	2.0×10^{-4}

Definition of abbreviations: BMI = body mass index, FE_{NO} = fractional exhaled nitric oxide. Data are presented as n (%), mean ± SD, or median (interquartile range).

Table 5. Expression Status of the Nine-Gene-Set Signature of Cluster A in Bronchial Biopsy and Epithelial Brushing

Group	Bronchial Biopsy	Epithelial Brushing	n (%)
Group 1	Expressed	Expressed	19 (23%)
Group 2	Expressed	Not expressed	17 (20%)
Group 3	Not expressed	Expressed	8 (10%)
Group 4	Not expressed	Not expressed	39 (47%)

greater level of oral steroid dependency would support the notion that relative corticosteroid insensitivity may be associated with corticosteroid dependency. A potential limitation of the analysis of relative corticosteroid insensitivity is the use of the gene signatures obtained from patients with COPD, in whom the gene response to corticosteroids may be different from that seen in patients with asthma.

Two gene sets used were related to ozone stimulation as a marker of elevated oxidative stress, and other important GSVa signatures that were critical in differentiating asthma included those from monocytes relating to autoimmune diseases and to Th1 immunity as reflected by the increase in the numbers of

submucosal CD3⁺ and CD8⁺ T cells. Among them, we noted that the oxidative stress signature revealed a moderate to high correlation with Th1 signature and with the steroid insensitivity signature (Figures E2 and E3). Although ROS has been reported to play a role in promoting asthmatic inflammation (14, 15) and reducing steroid responsiveness (16), the mechanism remains largely controversial. Therefore, the implication of oxidative stress in our findings may shed light on the impact of ROS in asthma. Overall, we found that there were non-Th2 pathways that were associated with severe asthma in the analysis of both bronchial biopsies and epithelial brushings, despite the presence of eosinophilic inflammation.

One potential concern regarding our approach in this study is the appropriateness of gene sets used for clustering. We endeavored to be as inclusive as possible by using gene sets reflecting the diversity of processes previously reported to be associated with various aspects of asthma pathobiology in addition to key mechanisms involved in chronic obstructive airway disease, baseline immune cell signatures, and those reported in a number of systemic immune disorders. In keeping with this deployment of gene sets, we found that group 1 subjects were distinguished by diverse immune activation across airway compartments, whereas nearly half of group 4 participants did not show pronounced activation of the immune cells selected. Interestingly, the enrichment of GSVa signatures representative of multiple immune pathways, including Th2, Th1, Th17, neutrophils, dendritic cells, and influenza response, was largely lower in group 4 than in group 1 (Figures E4 and E5), suggesting the lack of a dominant disease-driving immune pathway and possibly reflecting a pauciimmune phenotype of asthma. Further unbiased bioinformatic analyses, such as the use of weighted gene

Table 6. Clinical Features of the Expression of Nine Gene Set Signatures in Bronchial Biopsy and/or Epithelial Brushing

Variables	Group 1	Group 2	Group 3	Group 4	P Value		
					Group 1 vs. Group 4	Group 3 vs. Group 4	Group 1 vs. Group 3
n (%)	19 (23%)	17 (20%)	8 (10%)	39 (47%)			
Age, yr	48.0 ± 13.6	47.4 ± 14.1	50.6 ± 13.4	44.0 ± 13.5	0.284	0.187	0.644
Age of asthma onset, yr	24.7 ± 13.4	22.2 ± 14.4	32.0 ± 11.6	22.1 ± 13.7	0.485	0.039	0.176
Female sex	7 (36.8%)	8 (47.1%)	3 (37.5%)	24 (61.5%)	0.136	0.390	1.000
BMI, kg/m ²	27.8 ± 4.8	28.6 ± 4.9	33.1 ± 3.8	28.0 ± 5.9	0.986	0.010	0.007
Nasal polyps	8 (42.1%)	6 (35.3%)	2 (25.0%)	9 (23.1%)	0.102	0.898	0.357
Rhinitis	8 (42.1%)	8 (47.1%)	6 (75.0%)	16 (41.0%)	0.996	0.100	0.150
Eczema	6 (31.6%)	9 (52.9%)	2 (25.0%)	17 (43.6%)	0.604	0.525	0.900
Severe asthma	14 (73.7%)	10 (58.8%)	6 (75.0%)	20 (51.3%)	0.179	0.402	1.000
Oral corticosteroid use	9 (47.4%)	5 (29.4%)	2 (25.0%)	5 (12.8%)	0.010	0.737	0.515
Atopy	11 (57.9%)	12 (70.6%)	6 (75.0%)	30 (76.9%)	0.312	0.476	0.466
Exacerbations, yr ⁻¹	3.0 (2.0–5.0)	2.0 (1.3–2.8)	3.0 (2.0–3.0)	1.5 (1.0–3.0)	0.022	0.238	0.563
FEV ₁ , % predicted	68.1 (54.9–78.3)	90.4 (70.9–98.0)	80.3 (65.9–101.6)	79.7 (65.5–94.8)	0.133	0.708	0.260
Total IgE, IU/ml	163 (42–231.5)	86.6 (24.0–196.0)	93.9 (41.3–879.8)	96 (44.3–278.5)	0.659	0.839	0.815
Blood eosinophils, %	3.5 (2.4–8.0)	1.7 (1.4–3.6)	3.4 (3.2–4.2)	2.4 (1.4–3.9)	0.070	0.100	0.915
Blood neutrophils, %	58.8 (50.4–64.2)	61.9 (56.6–68.3)	65.5 (61.9–69.5)	58.6 (53.2–68.9)	0.596	0.380	0.217
Sputum eosinophils, %	8.2 (1.8–27.0)	0.6 (0.6–1.0)	15.8 (12.4–16.2)	0.6 (0–1.6)	0.017	0.014	0.863
Sputum neutrophils, %	52.3 (33.4–77.4)	57.0 (54.2–76.7)	45.3 (41.4–50.8)	52.7 (34.3–65.4)	0.835	0.763	0.727
FeNO, ppb	56.5 (34.1–74.6)	22.0 (14.5–32.5)	35.5 (20.0–50.4)	19 (12.5–36.1)	0.001	0.058	0.312
Submucosal cells, mm ⁻²							
Eosinophil	6.1 (4.3–13.6)	2.2 (0–5.9)	2.7 (1.5–4.6)	1.0 (0–3.5)	7.6E-06	0.109	0.019
Neutrophil	11.0 (4.6–18.2)	9.8 (7.1–18.7)	12.1 (6.5–13.5)	12.1 (6.8–19.2)	0.853	0.500	0.727
CD4 T cell	15.4 (9.2–24.7)	5.8 (1.5–11.1)	9.2 (6.2–16.3)	10.2 (4.6–15.5)	0.070	0.839	0.187
CD3 T cell	45.3 (32.7–77.2)	21.8 (16.0–37.0)	41.9 (36.9–70.9)	33.3 (16.4–49.9)	0.048	0.243	0.874
CD8 T cell	26.4 (15.7–37.0)	9.4 (4.6–14.2)	27.1 (18.1–39.8)	13.9 (4.7–22.1)	0.020	0.150	0.852

Definition of abbreviations: BMI = body mass index; FeNO = fractional exhaled nitric oxide. Data are presented as n (%), mean ± SD, or median (interquartile range).

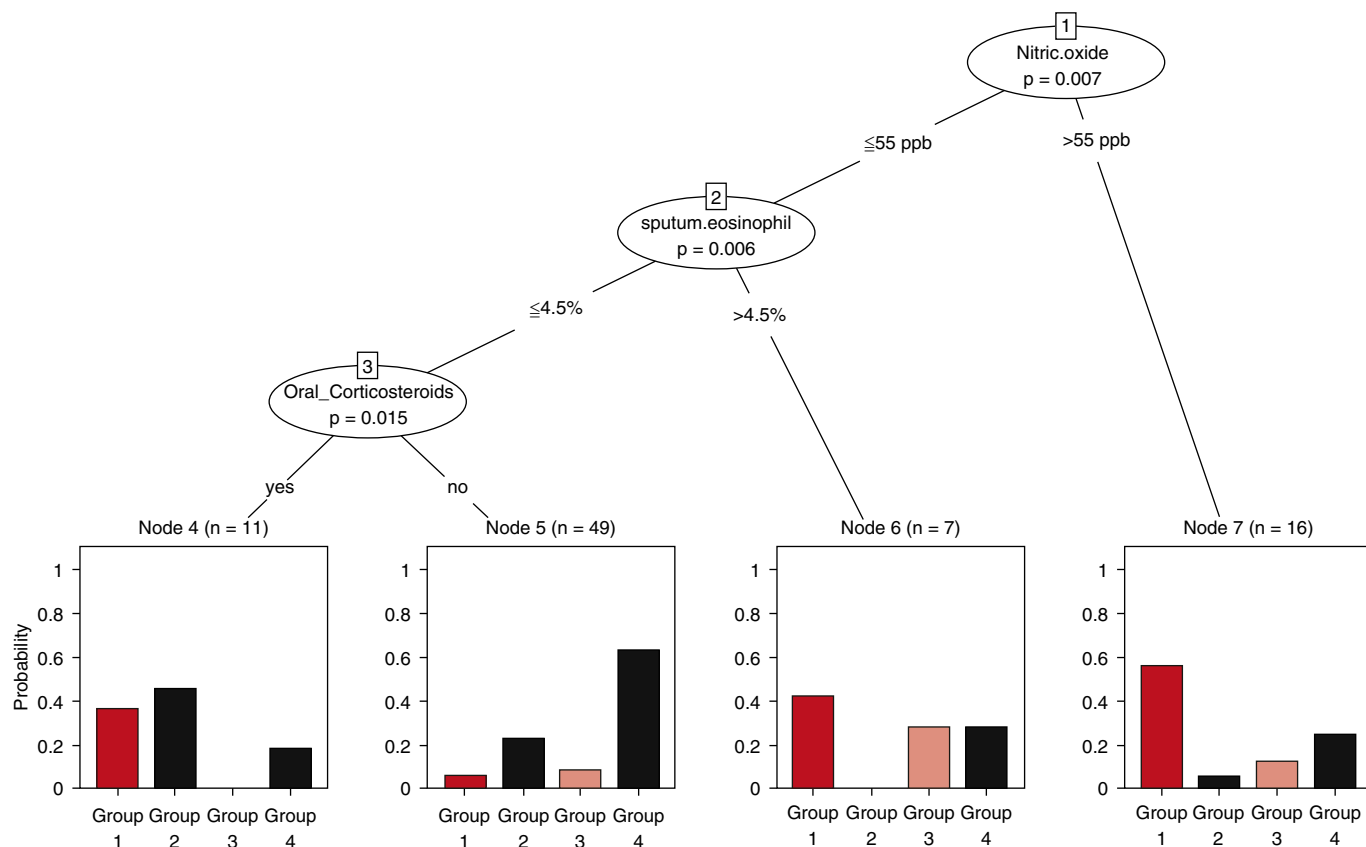


Figure 6. Tree-based inference scheme for the probability of group 1 to group 4 using two noninvasive measurements (fractional exhaled nitric oxide, $P = 0.014$; sputum eosinophil counts, $P = 0.012$) and one clinical parameter (oral corticosteroid dependency, $P = 0.030$). Group 1 (red bars) represented a phenotype of high eosinophilic inflammation and steroid insensitivity. Group 3 (pink bars) represented a phenotype of medium eosinophilic inflammation, and groups 2 and 4 (black bars) represented low eosinophilic inflammation.

coexpression network analysis, are necessary to determine other potential disease-driving mechanisms. In a preliminary weighted gene coexpression network analysis of epithelial cell gene expression, we found several modules related to FE_{NO} , sputum eosinophilia, and oral corticosteroid use with distinct eigengenes associated with these clinical parameters.

Although the GSVA nine-gene-set signature was initially derived from gene expression in bronchial biopsies, its expression in epithelial brushings was associated with similar clinical characteristics. The strength of this signature was greater when it was highly expressed in epithelial cells than when it was expressed solely in bronchial biopsies, particularly when linked to clinical parameters such as FE_{NO} , blood eosinophils, sputum eosinophils, and subtypes of submucosal T cells. This finding echoes previous studies highlighting the key role of

the bronchial epithelium in dictating Th2 asthma phenotype, and it emphasizes the importance of the bronchial epithelium in driving asthma pathobiology (42). However, cross-talk between these compartments must exist because the most severe asthma subtype we describe, featuring eosinophilic inflammation and corticosteroid insensitivity (group 1), occurred in subjects in whom the nine GSVA signatures were concomitantly highly expressed in both bronchial biopsies and epithelial brushings. Interestingly, the expression of the GSVA nine gene set signatures in epithelial cells was associated mainly with sputum eosinophilia (group 3 participants).

This observation may reflect priming of the asthmatic airway upon which the group 1 subtype was dependent when the signature was also found in the bronchial biopsies. This priming is further supported by the observation that subjects lacking this GSVA nine-gene-set signature expression in

epithelial cells do not show eosinophilic inflammation, regardless of their signature expression in bronchial biopsies (groups 2 and 4 participants). This highlights the need to further clarify the role played by airway epithelial cells in driving the development of airway inflammation in asthma. Indeed, recent work indicates that the airway epithelium is an important upstream cell that controls the regulation of Th2 cytokines through the production of cytokines such as TSLP, IL-25, and IL-33, which can be stimulated by exposure of epithelial cells to external stimuli such as pollutants, viruses, and allergens (42, 43). The epithelium in severe asthma is also reported to be thicker than in mild to moderate asthma, with increased proliferation, apoptosis, and release of proinflammatory factors (44). However, our results indicate that IL-33 and TSLP gene expression is increased in the biopsies, not in the brushings, of patients with asthma in cluster A and in group 1 who also show

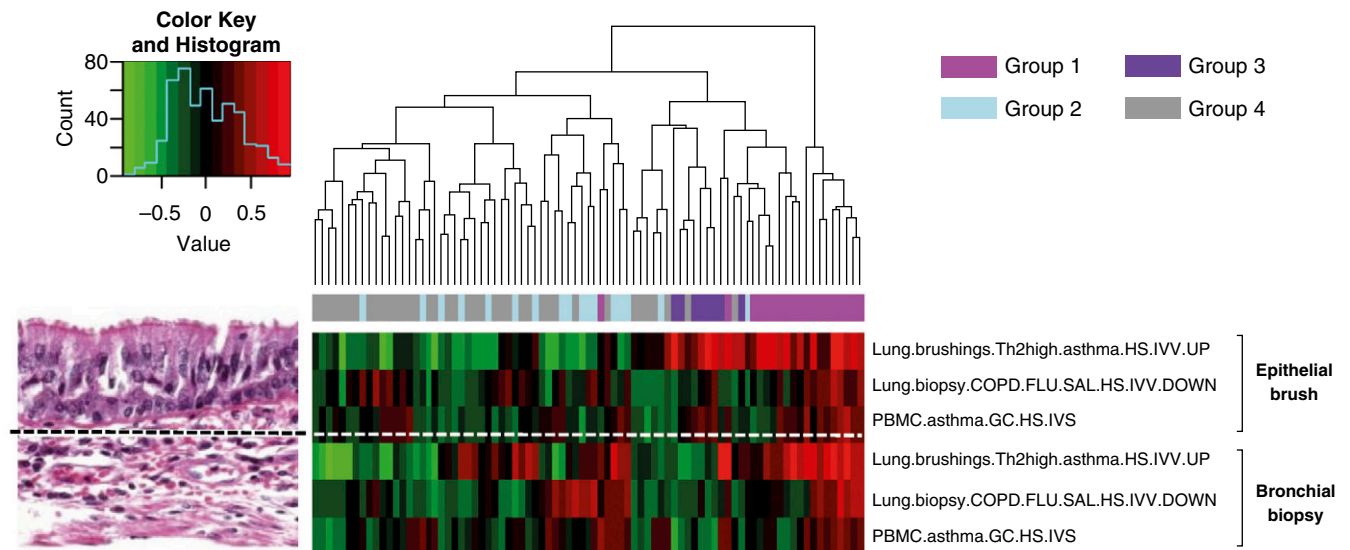


Figure 7. Three signatures related to T-helper cell type 2 (Th2) and steroid insensitivity in epithelial (*above the broken line*) and biopsy (*below the broken line*) cells across 83 subjects with asthma were used for clustering. Subject membership in groups 1–4 are shown as color bars and mapped underneath the dendrogram. Group 1 (*magenta*) patients with the highest and most extensive expression of these signatures were clustered mainly to the right. Group 3 (*purple*) subjects with the signatures solely in the epithelium also clustered closely. Groups 2 (*light blue*) and 4 (*gray*) subjects interlaced with each other and clustered to the left. These three out of the nine signatures were the major representative features of groups 1 and 3 asthma subtypes. COPD = chronic obstructive pulmonary disease; DOWN = down-regulated; FLU = fluticasone; GC = glucocorticoid; HS = *Homo sapiens*; IVS = *in vitro* soil-less; IVV = *in vivo*; PBMC = peripheral blood mononuclear cell; SAL = salmeterol; UP = up-regulated.

the highest Th2 gene signature with bronchial eosinophilia.

The subtypes of asthma reported here, based on transcriptomic rather than clinical data, share some overlapping clinical and biomarker features with phenotypes driven by clinical parameters from the Leicester (23) and Severe Asthma Research Program (SARP) (22) clustering studies. The small group 3 phenotype (10% of patients) reported here is very similar to cluster 3 in SARP (8%) in terms of higher body mass index and higher sputum eosinophils, but lower biopsy eosinophils, than in the remaining asthma groups. In addition, our group 1 phenotype (23%) has some resemblance to cluster 5 in SARP (16%) regarding the lower FEV₁, whereas these groups were discordant with respect to their neutrophilic profile. Indeed, we found no difference in the level of either blood or sputum neutrophilia in the four groups. These findings provide a bridge between phenotypes derived from clustering of clinical parameters and our approach of using gene expression profiles within key airway structural cells along with tissue histopathology. We have provided a path to exploring the mechanisms

underlying different subsets of patients in whom discordant inflammation is found.

A key question is whether there is any clinical utility or application of the phenotypes described here, particularly as they were determined using transcriptomic analysis of samples obtained by bronchoscopy. The clinical utility is addressed by the tree-based recursive partitioning algorithm selecting significant noninvasive parameters or clinical traits on the basis of their conditional distribution statistics against phenotypes using an unbiased permutation method (45). This approach successfully translated our phenotypes originally defined by gene signatures and by histopathologic results to a clinically inferable scheme. Following a stepwise binary value split of FE_{NO} (>55 ppb), sputum eosinophils (>4.5%), or oral corticosteroid dependency, this scheme identified group 1 subjects with an 84% sensitivity and 72% specificity. This is clinically important because this group bears the Th2 and steroid insensitivity signatures across airway structural cells that justifies the need for molecular agents that target Th2-specific pathobiology (20). However, other groups, such as groups 2 and 3, may

also benefit from these agents, with the possibility that different targets within the Th2 phenotype may be more appropriate for each of the three groups. Thus, epithelium-dependent release of different eosinophilic cytokines and/or chemokines might be appropriate for groups 1 and 3, whereas noneosinophilic Th2 or T2 targets might be more useful in Group 2.

In conclusion, transcriptomic profiling in bronchial biopsies and epithelial brushing cells showed heterogeneity of underlying mechanisms with respect to Th2 and relative steroid insensitivity signatures, which allowed the definition of different asthma phenotypes. Routinely used clinical biomarkers successfully translated these findings, and, used in this way, they may help clinicians in making therapeutic decisions for patients with severe eosinophilic inflammation and relative corticosteroid insensitivity. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank all the members of each recruiting center for their recruitment and assessment of the participants.

References

- Trejo Bittar HE, Yousem SA, Wenzel SE. Pathobiology of severe asthma. *Annu Rev Pathol* 2015;10:511–545.
- Bentley AM, Meng Q, Robinson DS, Hamid Q, Kay AB, Durham SR. Increases in activated T lymphocytes, eosinophils, and cytokine mRNA expression for interleukin-5 and granulocyte/macrophage colony-stimulating factor in bronchial biopsies after allergen inhalation challenge in atopic asthmatics. *Am J Respir Cell Mol Biol* 1993;8:35–42.
- Robinson DS, Ying S, Bentley AM, Meng Q, North J, Durham SR, Kay AB, Hamid Q. Relationships among numbers of bronchoalveolar lavage cells expressing messenger ribonucleic acid for cytokines, asthma symptoms, and airway methacholine responsiveness in atopic asthma. *J Allergy Clin Immunol* 1993;92:397–403.
- Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, Koth LL, Arron JR, Fahy JV. T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* 2009;180:388–395.
- Boniface S, Koscher V, Mamessier E, El Biaze M, Dupuy P, Lorec AM, Guillot C, Badier M, Bongrand P, Vervloet D, et al. Assessment of T lymphocyte cytokine production in induced sputum from asthmatics: a flow cytometry study. *Clin Exp Allergy* 2003;33:1238–1243.
- Cho SH, Stanciu LA, Holgate ST, Johnston SL. Increased interleukin-4, interleukin-5, and interferon- γ in airway CD4⁺ and CD8⁺ T cells in atopic asthma. *Am J Respir Crit Care Med* 2005;171:224–230.
- Krug N, Madden J, Redington AE, Lackie P, Djukanovic R, Schauer U, Holgate ST, Frew AJ, Howarth PH. T-cell cytokine profile evaluated at the single cell level in BAL and blood in allergic asthma. *Am J Respir Cell Mol Biol* 1996;14:319–326.
- Raundhal M, Morse C, Khare A, Oriss TB, Milosevic J, Trudeau J, Huff R, Pilewski J, Holguin F, Kolls J, et al. High IFN- γ and low SLPI mark severe asthma in mice and humans. *J Clin Invest* 2015;125:3037–3050.
- Chambers ES, Nanzer AM, Pfeffer PE, Richards DF, Timms PM, Martineau AR, Griffiths CJ, Corrigan CJ, Hawrylowicz CM. Distinct endotypes of steroid-resistant asthma characterized by IL-17A^{high} and IFN- γ ^{high} immunophenotypes: potential benefits of calcitriol. *J Allergy Clin Immunol* 2015;136:628–637.e4.
- McKinley L, Alcorn JF, Peterson A, Dupont RB, Kapadia S, Logar A, Henry A, Irvin CG, Piganelli JD, Ray A, et al. TH17 cells mediate steroid-resistant airway inflammation and airway hyperresponsiveness in mice. *J Immunol* 2008;181:4089–4097.
- Doe C, Bafadhel M, Siddiqui S, Desai D, Mistry V, Rugman P, McCormick M, Woods J, May R, Sleeman MA, et al. Expression of the T helper 17-associated cytokines IL-17A and IL-17F in asthma and COPD. *Chest* 2010;138:1140–1147.
- Comhair SA, Ricci KS, Aroliga M, Lara AR, Dweik RA, Song W, Hazen SL, Bleecker ER, Busse WW, Chung KF, et al. Correlation of systemic superoxide dismutase deficiency to airflow obstruction in asthma. *Am J Respir Crit Care Med* 2005;172:306–313.
- Kirkham P, Rahman I. Oxidative stress in asthma and COPD: antioxidants as a therapeutic strategy. *Pharmacol Ther* 2006;111:476–494.
- Comhair SA, Bhatena PR, Farver C, Thunnissen FB, Erzurum SC. Extracellular glutathione peroxidase induction in asthmatic lungs: evidence for redox regulation of expression in human airway epithelial cells. *FASEB J* 2001;15:70–78.
- Dweik RA, Comhair SA, Gaston B, Thunnissen FB, Farver C, Thomassen MJ, Kavuru M, Hammel J, Abu-Soud HM, Erzurum SC. NO chemical events in the human airway during the immediate and late antigen-induced asthmatic response. *Proc Natl Acad Sci USA* 2001;98:2622–2627.
- Ito K, Hanazawa T, Tomita K, Barnes PJ, Adcock IM. Oxidative stress reduces histone deacetylase 2 activity and enhances IL-8 gene expression: role of tyrosine nitration. *Biochem Biophys Res Commun* 2004;315:240–245.
- James AL, Bai TR, Mauad T, Abramson MJ, Dolnikoff M, McKay KO, Maxwell PS, Elliot JG, Green FH. Airway smooth muscle thickness in asthma is related to severity but not duration of asthma. *Eur Respir J* 2009;34:1040–1045.
- Kaminska M, Foley S, Maghni K, Storness-Bliss C, Coxson H, Ghezzi H, Lemièr C, Olivenstein R, Ernst P, Hamid Q, et al. Airway remodeling in subjects with severe asthma with or without chronic persistent airflow obstruction. *J Allergy Clin Immunol* 2009;124:45–51.e4.
- Macedo P, Hew M, Torrego A, Jounneau S, Oates T, Durham A, Chung KF. Inflammatory biomarkers in airways of patients with severe asthma compared with non-severe asthma. *Clin Exp Allergy* 2009;39:1668–1676.
- Chung KF. New treatments for severe treatment-resistant asthma: targeting the right patient. *Lancet Respir Med* 2013;1:639–652.
- Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* 2012;18:716–725.
- Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, Li X, D'Agostino R Jr, Castro M, Curran-Everett D, Fitzpatrick AM, et al.; National Heart, Lung, and Blood Institute's Severe Asthma Research Program. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2010;181:315–323.
- Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, Wardlaw AJ, Green RH. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 2008;178:218–224.
- Bel EH, Sousa A, Fleming L, Bush A, Chung KF, Versnel J, Wagener AH, Wagers SS, Sterk PJ, Compton CH; Unbiased Biomarkers for the Prediction of Respiratory Disease Outcome (U-BIOPRED) Consortium, Consensus Generation. Diagnosis and definition of severe refractory asthma: an international consensus statement from the Innovative Medicine Initiative (IMI). *Thorax* 2011;66:910–917.
- Miller MK, Johnson C, Miller DP, Deniz Y, Bleecker ER, Wenzel SE; TENOR Study Group. Severity assessment in asthma: an evolving concept. *J Allergy Clin Immunol* 2005;116:990–995.
- Wenzel SE. Asthma: defining of the persistent adult phenotypes. *Lancet* 2006;368:804–813.
- Chung KF, Adcock IM. Clinical phenotypes of asthma should link up with disease mechanisms. *Curr Opin Allergy Clin Immunol* 2015;15:56–62.
- Pavord ID, Korn S, Howarth P, Bleecker ER, Buhl R, Keene ON, Ortega H, Chanez P. Mepolizumab for severe eosinophilic asthma (DREAM): a multicentre, double-blind, placebo-controlled trial. *Lancet* 2012;380:651–659.
- Chung KF. Defining phenotypes in asthma: a step towards personalized medicine. *Drugs* 2014;74:719–728.
- Yan X, Chu JH, Gomez J, Koenigs M, Holm C, He X, Perez MF, Zhao H, Mane S, Martinez FD, et al. Noninvasive analysis of the sputum transcriptome discriminates clinical phenotypes of asthma. *Am J Respir Crit Care Med* 2015;191:1116–1125.
- Edelman E, Porrello A, Guinney J, Balakumaran B, Bild A, Febbo PG, Mukherjee S. Analysis of sample set enrichment scores: assaying the enrichment of sets of genes for individual samples in genome-wide expression profiles. *Bioinformatics* 2006;22:e108–e116.
- Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
- Bao Z, Zhang C, Yan W, Liu Y, Li M, Zhang W, Jiang T. BMP4, a strong better prognosis predictor, has a subtype preference and cell development association in gliomas. *J Transl Med* 2013;11:100.
- Nicolau M, Levine AJ, Carlsson G. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proc Natl Acad Sci USA* 2011;108:7265–7270.
- Sarikonda G, Pettus J, Phatak S, Sachithanatham S, Miller JF, Wesley JD, Cadag E, Chae J, Ganesan L, Mallios R, et al. CD8 T-cell reactivity to islet antigens is unique to type 1 while CD4 T-cell reactivity exists in both type 1 and type 2 diabetes. *J Autoimmun* 2014;50:77–82.
- Thalamuthu A, Mukhopadhyay I, Zheng X, Tseng GC. Evaluation and comparison of gene clustering methods in microarray analysis. *Bioinformatics* 2006;22:2405–2412.
- Auffray C, Adcock IM, Chung KF, Djukanovic R, Pison C, Sterk PJ. An integrative systems biology approach to understanding pulmonary diseases. *Chest* 2010;137:1410–1416.
- Kuo CHS, Pavlidis S, Loza M, Baribaud F, Rowe A, Pandis I, Rossios C, Wilson SJ, Djukanovic R, Sterk PJ, et al. Asthma phenotypes from a semi-supervised machine-learning approach of bronchial biopsy and

- brush transcriptomics in U-BIOPRED [abstract]. *Am J Respir Crit Care Med* 2015;191:A2205.
39. Shaw DE, Sousa AR, Fowler SJ, Fleming LJ, Roberts G, Corfield J, Pandis I, Bansal AT, Bel EH, Auffray C, *et al.* Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J* 2015;46:1308–1321.
 40. Atthey BD, Braxenthaler M, Haas M, Guo Y. tranSMART: an open source and community-driven informatics and data sharing platform for clinical and translational research. *AMIA Jt Summits Transl Sci Proc* 2013;2013:6–8.
 41. Abbas AR, Jackman JK, Bullens SL, Davis SM, Choy DF, Fedorowicz G, Tan M, Truong BT, Gloria Meng Y, Diehl L, *et al.* Lung gene expression in a rhesus allergic asthma model correlates with physiologic parameters of disease and exhibits common and distinct pathways with human asthma and a mouse asthma model. *Am J Pathol* 2011;179:1667–1680.
 42. Holgate ST. The sentinel role of the airway epithelium in asthma pathogenesis. *Immunol Rev* 2011;242:205–219.
 43. Licona-Limón P, Kim LK, Palm NW, Flavell RA. TH2, allergy and group 2 innate lymphoid cells. *Nat Immunol* 2013;14:536–542.
 44. Cohen L, E X, Tarsi J, Ramkumar T, Horiuchi TK, Cochran R, DeMartino S, Schechtman KB, Hussain I, Holtzman MJ, *et al.*; NHLBI Severe Asthma Research Program (SARP). Epithelial cell proliferation contributes to airway remodeling in severe asthma. *Am J Respir Crit Care Med* 2007;176:138–145.
 45. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. *J Comput Graph Stat* 2006;15:651–674.